1. Introduction

In this paper we take up the analysis of dynamic simultaneous equation models (SEM's) within the context of general linear multiple time series processes such as studied by Quenouille (1957). As noted by Quenouille, if a set of variables is generated by a multiple time series process, it is often possible to solve for the processes generating individual variables, namely the 'final equations' of Tinbergen (1940), and these are in the autoregressive-moving average (ARMA) form. ARMA processes have been studied intensively by Box and Jenkins (1970). Further, if a general multiple time series process is appropriately specialized, we obtain a usual dynamic SEM in structural form. By algebraic manipulations, the associated reduced form and transfer function equation systems can be derived. In what follows, these equation systems are presented and their properties and uses are indicated.

It will be shown that assumptions about variables being exogenous, about lags in structural equations of SEM's, and about serial correlation properties of structural disturbance terms have strong implications for the properties of transfer functions and final equations that can be tested. Further, we show how large sample posterior odds and likelihood ratios can be used to appraise alternative hypotheses. In agreement with Pierce and Mason (1971), we believe that testing the implications of structural assumptions for transfer functions and, we add, final equations is an important element in the process of iterating in on a model that is reasonably in accord with the information in a sample of data. To illustrate these general points and to provide applications of the above methods, a
dynamic version of a SEM due to Haavelmo (1947) is analyzed using U.S. post-
World War II quarterly data.

The plan of the paper is as follows. In sect. 2, a general multiple time series
model is specified, its final equations are obtained and their properties set forth.
Then the implications of assumptions needed to specialize the multiple time series
model to become a dynamic SEM for transfer functions and final equations are
presented. In sect. 3, the algebraic analysis is applied to a small dynamic SEM.
Quarterly U.S. data are employed in sects. 4 and 5 to analyze the final and transfer
equations of the dynamic SEM. Sect. 6 provides a discussion of the empirical
results, their implications for the specification and estimation of the structural
equations of the model, and some concluding remarks.

2. General formulation and analysis of a system of dynamic equations

As indicated by Quenouille (1957), a linear multiple time series process can be
represented as follows:¹

\[ H(L) z_t = F(L) e_t, \quad t = 1, 2, \ldots, T, \] (2.1)

where \( z_t = (z_{1t}, z_{2t}, \ldots, z_{pt}) \) is a vector of random variables, \( e_t = (e_{1t}, e_{2t}, \ldots, e_{pt}) \) is a vector of random errors, and \( H(L) \) and \( F(L) \) are each \( p \times p \) matrices,
assumed of full rank, whose elements are finite polynomials in the lag operator
\( L \), defined as \( L z_t = z_{t-1} \). Typical elements of \( H(L) \) and \( F(L) \) are given by 

\[ h_{ij} = \sum_{l=0}^{r} h_{ijl} L^l \] and 

\[ f_{ij} = \sum_{l=0}^{q} f_{ijl} L^l. \] Further, we assume that the error process has
a zero mean, an identity covariance matrix and no serial correlation, that is:

\[ E e_t = 0, \quad \text{for all } t \text{ and } t', \] (2.2)

\[ E e_t e_{t'} = \delta_{tt'} I, \] (2.3)

where \( I \) is a unit matrix and \( \delta_{tt'} \) is the Kronecker delta. The assumption in (2.3)
does not involve a loss of generality since correlation of errors can be introduced
through the matrix \( F(L) \).

The model in (2.1) is a multivariate autoregressive-moving average (ARMA)
process. If \( H(L) = H_0 \), a matrix of degree zero in \( L \), (2.1) is a moving average
(MA) process; if \( F(L) = F_0 \), a matrix of degree zero in \( L \), it is an autoregressive
(AR) process. In general, (2.1) can be expressed as:

\[ \sum_{l=0}^{r} H_l L^l z_t = \sum_{l=0}^{q} F_l L^l e_t, \] (2.4)

where \( H_l \) and \( F_l \) are matrices with all elements not depending on \( L \), \( r = \max_{i,j} r_{ij} \) and \( q = \max_{i,j} q_{ij} \).

¹In (2.1), \( z_t \) is assumed to be mean-corrected, that is \( z_t \) is a deviation from a population mean
vector. Below we relax this assumption.
Since $H(L)$ in (2.1) is assumed to have full rank, (2.1) can be solved for $z_t$ as follows:

$$z_t = H^{-1}(L)F(L)e_t,$$
(2.5a)

or

$$z_t = [H^*(L)/|H(L)|]F(L)e_t,$$
(2.5b)

where $H^*(L)$ is the adjoint matrix associated with $H(L)$ and $|H(L)|$ is the determinant which is a scalar, finite polynomial in $L$. If the process is to be invertible, the roots of $|H(L)| = 0$ have to lie outside the unit circle. Then (2.5) expresses $z_t$ as an infinite MA process that can be equivalently expressed as the following system of finite order ARMA equations:

$$|H(L)|z_t = H^*(L)F(L)e_t.$$  
(2.6)

The $i$th equation of (2.6) is given by:

$$|H(L)|z_{it} = \alpha'_i e_t, \quad i = 1, 2, \ldots, p,$$
(2.7)

where $\alpha'_i$ is the $i$th row of $H^*(L)F(L)$.

The following points regarding the set of final equations in (2.7) are of interest:

(i) Each equation is in ARMA form, as pointed out by Quenouille (1957, p. 20). Thus the ARMA processes for individual variables are compatible with some, perhaps unknown, joint process for a set of random variables and are thus not necessarily 'naive', 'ad hoc' alternative models.

(ii) The order and parameters of the autoregressive part of each equation, $|H(L)|z_{it}$, $i = 1, 2, \ldots, p$, will usually be the same.²

(iii) Statistical methods can be employed to investigate the form and properties of the ARMA equations in (2.7). Given that their forms, that is the degree of $|H(L)|$ and the order of the moving average errors, have been determined, they can be estimated and used for prediction.

(iv) The equations of (2.7) are in the form of a restricted 'seemingly unrelated' autoregressive model with correlated moving average error processes.³

The general multiple time series model in (2.1) can be specialized to a usual dynamic simultaneous equation model (SEM) if some prior information about $H$ and $F$ is available. That is, prior information may indicate that it is appropriate to regard some of the variables in $z_t$ as being endogenous and the remaining variables as being exogenous, that is, generated by an independent process. To represent this situation, we partition (2.1) as follows:

$$
\begin{pmatrix}
H_{11} & H_{12} \\
H_{21} & H_{22}
\end{pmatrix}
\begin{pmatrix}
y_t \\
x_t
\end{pmatrix}
= 
\begin{pmatrix}
F_{11} & F_{12} \\
F_{21} & F_{22}
\end{pmatrix}
\begin{pmatrix}
e_{1t} \\
e_{2t}
\end{pmatrix}. 
$$
(2.8)

²In some cases in which $|H(L)|$ contains factors in common with those appearing in all elements of the vectors $\alpha'_i$, e.g. when $H$ is triangular, diagonal or block diagonal, some canceling will take place. In such cases the statement in (ii) has to be qualified.

³See Nelson (1970) and Akaike (1973) for estimation results for systems similar to (2.7).
If the $p_1 \times 1$ vector $y_t$ is endogenous and the $p_2 \times 1$ vector $x_t$ is exogenous, this implies the following restrictions on the submatrices of $H$ and $F$:

$$H_{21} \equiv 0, \quad F_{21} \equiv 0, \quad \text{and} \quad F_{12} \equiv 0. \quad (2.9)$$

With the assumptions in (2.9), the elements of $e_{1t}$ do not affect the elements of $x_t$ and the elements of $e_{2t}$ affect the elements of $y_t$ only through the elements of $x_t$. Under the hypotheses in (2.9), (2.8) is in the form of a dynamic SEM with endogenous variable vector $y_t$ and exogenous variable vector $x_t$ generated by an ARMA process. The usual structural equations, from (2.8) subject to (2.9), are:

$$H_{11}(L) y_t + H_{12}(L) x_t = F_{11}(L) e_{1t}, \quad (2.10)$$

while the process generating the exogenous variables is:

$$H_{22}(L) x_t = F_{22}(L) e_{2t}, \quad (2.11)$$

with $p_1 + p_2 = p$.

Analogous to (2.4), the system (2.10) can be expressed as:

$$\sum_{l=0}^{r} H_{11} L^l y_t + \sum_{l=0}^{r} H_{12} L^l x_t = \sum_{l=0}^{r} F_{11} L^l e_{1t}, \quad (2.12)$$

where $H_{11}$, $H_{12}$, and $F_{11}$ are matrices the elements of which are coefficients of $L^i$. Under the assumption that $H_{110}$ is of full rank, the reduced form equations, which express the current values of endogenous variables as functions of the lagged endogenous and current and lagged exogenous variables, are:

$$y_t = - \sum_{l=0}^{r} H_{110}^{-1} H_{11} L^l y_t - \sum_{l=0}^{r} H_{110}^{-1} H_{12} L^l x_t$$

$$+ \sum_{l=0}^{q} H_{110}^{-1} F_{11} L^l e_{1t}. \quad (2.13)$$

The reduced form system in (2.13) is a system of $p_1$ stochastic difference equations of maximal order $r$.

The ‘final form’ of (2.13), Theil and Boot (1962), or ‘set of fundamental dynamic equations’ associated with (2.13), Kmenta (1971), which expresses the current values of endogenous variables as functions of only the exogenous variables, is given by:

$$y_t = - H_{11}^{-1}(L) H_{12}(L) x_t + H_{11}^{-1}(L) F_{11}(L) e_{1t}. \quad (2.14)$$

If the process is invertible, i.e. if the roots of $|H_{11}(L)| = 0$ lie outside the unit circle, (2.14) is an infinite MA process in $x_t$ and $e_{1t}$. Note that (2.14) is a set of ‘rational distributed lag’ equations, Jorgenson (1966), or a system of ‘transfer

4 Hannan (1969, 1971) has analyzed the identification problem for systems in the form of (2.10).
function' equations, Box and Jenkins (1970). Also, the system in (2.14) can be brought into the following form:

\[ |H_{11}(L)|y_t = -H_{11}^*(L)H_{12}(L)x_t + H_{11}^*(L)F_{11}(L)e_{11}, \]  

(2.15)

where \( H_{11}^*(L) \) is the adjoint matrix associated with \( H_{11}(L) \) and \( |H_{11}(L)| \) is the determinant of \( H_{11}(L) \). The equation system in (2.15), where each endogenous variable depends only on its own lagged values and on the exogenous variables, with or without lags, has been called the 'separated form', Marschak (1950), 'autoregressive final form', Dhrymes (1970), 'transfer function form', Box and Jenkins (1970), or 'fundamental dynamic equations', Pierce and Mason (1971).

As in (2.7), the \( p_1 \) endogenous variables in \( y_t \) have autoregressive parts with identical order and parameters, a point emphasized by Pierce and Mason (1971).

Having presented several equation systems above, it is useful to consider their possible uses and some requirements that must be met for these uses. As noted above, the final equations in (2.7) can be used to predict the future values of some or all variables in \( z_t \), given that the forms of the ARMA processes for these variables have been determined and that parameters have been estimated. However, these final equations cannot be used for control and structural analysis. On the other hand, the reduced form equations (2.13) and transfer equations (2.15) can be employed for both prediction and control but not generally for structural analysis except when structural equations are in reduced form \([H_{110} = I \text{ in } (2.12)]\) or in final form \([H_{11} = I \text{ in } (2.10)]\). Note that use of reduced form and transfer function equations implies that we have enough prior information to distinguish endogenous and exogenous variables. Further, if data on some of the endogenous variables are unavailable, it may be impossible to use the reduced form equations whereas it will be possible to use the transfer equations relating to those endogenous variables for which data are available. When the structural equation system in (2.10) is available, it can be employed for structural analysis and the associated 'restricted' reduced form or transfer equations can be employed for prediction and control. Use of the structural system (2.10) implies not only that endogenous and exogenous variables have been distinguished, but also that prior information is available to identify structural parameters and that the dynamic properties of the structural equations have been determined. Also, structural analysis of the complete system in (2.10) will usually require that data be available on all variables. For the reader's convenience, some of these considerations are summarized in table 1.

Aside from the differing data requirements for use of the various equation systems considered in table 1, it should be appreciated that before each of the equation systems can be employed, the form of its equations must be ascertained. For example, in the case of the structural equation system (2.10), not only must

5If some of the variables in \( x \) are non-stochastic, say time trends, they will appear the final equations of the system.

6This requirement will not be as stringent for partial analyses and for fully recursive models.
endogenous and exogenous variables be distinguished, but also lag distributions, serial correlation properties of error terms, and identifying restrictions must be specified. Since these are often difficult requirements, it may be that some of the simpler equation systems will often be used although their uses are more limited than those of structural equation systems. Furthermore, even when the objective

Table 1
Uses and requirements for various equation systems.

<table>
<thead>
<tr>
<th>Equation system</th>
<th>Uses of equation systems</th>
<th>Requirements for use of equation systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Final equations* (2.7)</td>
<td>yes no no</td>
<td>Forms of ARMA processes and parameter estimates</td>
</tr>
<tr>
<td>2. Reduced form equations (2.13)</td>
<td>yes yes no</td>
<td>Endogenous–exogenous classification of variables, forms of equations, and parameter estimates</td>
</tr>
<tr>
<td>3. Transfer equationsb (2.15)</td>
<td>yes yes no</td>
<td>Endogenous–exogenous classification of variables, forms of equations, and parameter estimates</td>
</tr>
<tr>
<td>4. Final form equationsc (2.14)</td>
<td>yes yes no</td>
<td>Endogenous–exogenous classification of variables, forms of equations, and parameter estimates</td>
</tr>
<tr>
<td>5. Structural equations (2.10)</td>
<td>yes yes yes</td>
<td>Endogenous–exogenous variable classification, identifying information, forms of equations, and parameter estimates</td>
</tr>
</tbody>
</table>

*aThis is Tinbergen’s (1940) term.
bThese equations are also referred to as ‘separated form’ or ‘autoregressive final form’ equations.
cAs noted in the text, these equations are also referred to as ‘transfer function’, ‘fundamental dynamic’, and ‘rational distributed lag’ equations.
dThat is, information in the form of restrictions to identify structural parameters.

of an analysis is to obtain a structural equation system, the other equation systems, particularly the final equations and transfer equations, will be found useful. That is, structural assumptions regarding lag structures, etc. have implications for the forms and properties of final and transfer equations that can be checked with data. Such checks on structural assumptions can reveal weaknesses
in them and possibly suggest alternative structural assumptions more in accord with the information in the data. In the following sections we illustrate these points in the analysis of a small dynamic structural equation system.

3. Algebraic analysis of a dynamic version of Haavelmo's model

Haavelmo (1947) formulated and analyzed the following static model with annual data for the U.S., 1929–1941:

\[ \begin{align*}
  c_t &= \alpha y_t + \beta + u_t, \\
  r_t &= \mu (c_t + x_t) + v + w_t, \\
  y_t &= c_t + x_t - r_t
\end{align*} \]  

(3.1)

where \( c_t, y_t \) and \( r_t \) are endogenous variables, \( x_t \) is exogenous, \( u_t \) and \( w_t \) are disturbance terms, and \( \alpha, \beta, \mu \) and \( v \) are scalar parameters. The definitions of the variables, all on a price-deflated, per capita basis, are:

\( c_t \) = personal consumption expenditures,

\( y_t \) = personal disposable income,

\( r_t \) = gross business saving, and

\( x_t \) = gross investment.\(^7\)

Eq. (3.1a) is a consumption relation, (3.1b) a gross business saving equation and (3.1c) an accounting identity.

In Chetty's (1966, 1968) analyses of the system (3.1) employing Haavelmo's annual data, he found the disturbance terms highly autocorrelated, perhaps indicating that the static nature of the model is not appropriate. In view of this possibility, (3.1) is made dynamic in the following way:

\[ \begin{align*}
  c_t &= \alpha(L) y_t + \beta + u_t, \\
  r_t &= \mu(L)(c_t + x_t) + v + w_t, \\
  y_t &= c_t + x_t - r_t
\end{align*} \]  

(3.2)

In (3.2a), \( \alpha(L) \) is a polynomial lag operator that serves to make \( c_t \) a function of current and lagged values of income. Similarly, \( \mu(L) \) in (3.2b) is a polynomial lag

\(^7\)In Haavelmo's paper, gross investment, \( x_t \), is defined equal to 'government expenditures + transfers - all taxes + gross private capital formation', while gross business saving, \( r_t \), is defined equal to 'depreciation and depletion charges + capital outlay charged to current expense + income credited to other business reserves - revaluation of business inventories + corporate savings'.

operator that makes \( r_t \) depend on current and lagged values of \( c_t + x_t \), a variable that Haavelmo refers to as 'gross disposable income'. On substituting for \( r_t \) in (3.2b) from (3.2c), the equations for \( c_t \) and \( y_t \) are:

\[
c_t = \alpha(L)y_t + \beta + u_t, \tag{3.3a}
\]

\[
y_t = [1 - \mu(L)](c_t + x_t) - \nu - w_t. \tag{3.3b}
\]

With respect to the disturbance terms in (3.3), we assume:

\[
\begin{pmatrix}
  u_t \\
  -w_t
\end{pmatrix}
= \begin{pmatrix}
  f_{11}(L) & f_{12}(L) & e_{1t} \\
  f_{21}(L) & f_{22}(L) & e_{2t}
\end{pmatrix},
\tag{3.4}
\]

where the \( f_{ij}(L) \) are polynomials in \( L \), \( e_{1t} \) and \( e_{2t} \) have zero means, unit variances, and are contemporaneously and serially uncorrelated.

Letting \( z_t' = (c_t, y_t, x_t) \), the general multiple time series model for \( z_t \), in the matrix form (2.1), is:

\[
H(L) z_t = \theta + F(L) e_t, \tag{3.5}
\]

where \( e_t' = (e_{1t}, e_{2t}, e_{3t}) \) satisfies (2.2)–(2.3) and \( \theta' = (\theta_1, \theta_2, \theta_3) \) is a vector of constants. In explicit form, (3.5) is:

\[
\begin{bmatrix}
  h_{11}(L) & h_{12}(L) & h_{13}(L) \\
  h_{21}(L) & h_{22}(L) & h_{23}(L) \\
  h_{31}(L) & h_{32}(L) & h_{33}(L)
\end{bmatrix}
\begin{bmatrix}
  c_t \\
  y_t \\
  x_t
\end{bmatrix}
= \begin{bmatrix}
  \theta_1 \\
  \theta_2 \\
  \theta_3
\end{bmatrix}
+ \begin{bmatrix}
  f_{11}(L) & f_{12}(L) & f_{13}(L) \\
  f_{21}(L) & f_{22}(L) & f_{23}(L) \\
  f_{31}(L) & f_{32}(L) & f_{33}(L)
\end{bmatrix}
\begin{bmatrix}
  e_{1t} \\
  e_{2t} \\
  e_{3t}
\end{bmatrix}. \tag{3.6}
\]

To specialize (3.6) to represent the dynamic version of Haavelmo's model in (3.3) with \( x_t \) exogenous, we must have \( \theta_1 = \beta, \theta_2 = \nu, \)

\[
\begin{align*}
  h_{11}(L) & \equiv 1 & h_{12}(L) & \equiv -\alpha(L) & h_{13}(L) & \equiv 0 \\
  h_{21}(L) & \equiv -[1 - \mu(L)] & h_{22}(L) & \equiv 1 & h_{23}(L) & \equiv -[1 - \mu(L)] \\
  h_{31}(L) & \equiv 0 & h_{32}(L) & \equiv 0 & h_{33}(L) & \equiv 0
\end{align*} \tag{3.7a}
\]

and

\[
\begin{align*}
  f_{13}(L) & \equiv f_{23}(L) \equiv f_{31}(L) \equiv f_{32}(L) \equiv 0.
\end{align*} \tag{3.7b}
\]
Utilizing the conditions in (3.7), (3.6) becomes:

\[
\begin{pmatrix}
1 & h_{12}(L) & 0 \\
0 & 0 & h_{21}(L)
\end{pmatrix}
\begin{pmatrix}
1 & h_{12}(L) & 0 \\
0 & 0 & h_{21}(L)
\end{pmatrix}
\begin{pmatrix}
c_t \\
y_t \\
x_t
\end{pmatrix}
= \begin{pmatrix}
\theta_1 \\
\theta_2 \\
\theta_3
\end{pmatrix}
+ \begin{pmatrix}
f_{11}(L) & f_{12}(L) & 0 \\
f_{21}(L) & f_{22}(L) & 0
\end{pmatrix}
\begin{pmatrix}
e_{1t} \\
e_{2t} \\
e_{3t}
\end{pmatrix}.
\]

(3.8)

Note that the process on the exogenous variable is \(h_{33}(L)x_t = f_{33}(L)e_{3t} + \theta_3\) and the fact that \(x_t\) is assumed exogenous requires that \(h_{31}(L) = h_{32}(L) = 0\) and that \(F(L)\) be block diagonal as shown in (3.8).

### Table 2

Degrees of lag polynomials in final equations.

<table>
<thead>
<tr>
<th>Final equation</th>
<th>Degrees of AR polynomials*</th>
<th>Degrees of MA polynomials for errorsb</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3.9): (c_t)</td>
<td>Maximum of (r_{33}) and (r_{12} + r_{21} + r_{33})</td>
<td>Maximum of (r_{33} + q_{11}) and (r_{33} + q_{22}) and (r_{12} + r_{23} + q_{33})</td>
</tr>
<tr>
<td>(3.10): (y_t)</td>
<td>(r_{33}) and (r_{12} + r_{21} + r_{33})</td>
<td>(r_{33} + q_{21}) and (r_{33} + q_{33}) and (r_{23} + q_{33})</td>
</tr>
<tr>
<td>(3.11): (x_t)</td>
<td>(r_{33})</td>
<td>(r_{33} + q_{11}) and (r_{33} + q_{22}) and (r_{23} + q_{33})</td>
</tr>
</tbody>
</table>

*a\(r_{ij}\) is the degree of \(h_{ij}\). Note from (3.7a), \(h_{21} = h_{23} = [1 - \mu(L)]\), and thus \(r_{21} - r_{23}\).

b\(q_{ij}\) is the degree of \(f_{ij}\).

In what follows we shall denote the degree of \(h_{ij}(L)\) by \(r_{ij}\) and the degree of \(f_{ij}(L)\) by \(q_{ij}\).

From (3.8), the final equations for \(c_t\) and \(y_t\) are given by:

\[
(1 - h_{12} h_{21}) h_{33} c_t = \theta'_1 + (f_{11} - f_{21} h_{12}) h_{33} e_{1t},
\]

(3.9)

\[
+ (f_{12} - f_{22} h_{12}) h_{33} e_{2t} + f_{33} h_{12} h_{23} e_{3t},
\]

and

\[
(1 - h_{12} h_{21}) h_{33} y_t = \theta'_2 + (f_{21} - f_{11} h_{21}) h_{33} e_{1t},
\]

(3.10)

\[
+ (f_{22} - f_{12} h_{21}) h_{33} e_{2t} - f_{33} h_{12} h_{23} e_{3t},
\]

with \(h_{21} \equiv h_{23}\) and \(\theta'_1\) and \(\theta'_2\) being new constants. Note that the AR parts of (3.9) and (3.10) have the same order and parameters. The degrees of the lag
polynomials in (3.9) and (3.10) and in the process for \( x_t \),

\[
h_{33}x_t = f_{33}e_{3t} + \theta_3,
\]

(3.11)

are indicated in table 2.

As mentioned above, the AR polynomials in the final equations for \( c_t \) and \( y_t \) are identical and of maximal degree equal to \( r_{12} + r_{21} + r_{33} \), as shown in table 2, where \( r_{12} \) is the degree of \( f(L) \) in the consumption equation, \( r_{21} \) is the degree of \( \mu(L) \) in the business saving equation, and \( r_{33} \) is the degree of \( h_{33} \), the AR polynomial in the process for \( x_t \). Also, if the disturbance terms \( u_t \) and \( w_t \) are serially uncorrelated and if all the \( q_{ij} \) in table 2 are zero, the following results hold:

(i) In the final equation for \( c_t \), the degree of the AR part is larger than or equal to the order of the MA process for the disturbance term; that is \( r_{12} + r_{21} + r_{33} \geq \max(r_{12} + r_{23}, r_{33} + r_{12}) \), with equality holding if \( r_{33} = 0 \), since \( r_{21} = r_{23} \), or if \( r_{12} = r_{23} = 0 \).

(ii) In the final equation for \( y_t \), the degree of the AR polynomial is larger than or equal to the order of the MA process for the disturbance term; i.e., \( r_{12} + r_{21} + r_{33} \geq r_{33} + r_{21} \) with equality holding if \( r_{12} = 0 \). Thus if the process for \( x_t \) is purely AR and the structural disturbance terms \( u_t \) and \( w_t \) are not serially correlated, (i) and (ii) provide useful implications for properties of the final equations that can be checked with data as explained below.

Further, under the assumption that the structural disturbance terms \( u_t \) and \( w_t \) are serially uncorrelated, all \( q_{ij} \) other than \( q_{33} \) in table 2 will be equal to zero. If the process for \( x_t \) is analyzed to determine the degree of \( h_{33} \), \( r_{33} \), and of \( f_{33}, q_{33} \), this information can be used in conjunction with the following:

(iii) In the final equation for \( c_t \), the degree of the AR polynomial will be smaller than or equal to the order of the MA disturbance if \( q_{33} \geq r_{33} \). (Note \( r_{21} = r_{23} \).) If \( q_{33} < r_{33} \), the degree of the AR polynomial will be greater than the order of the MA disturbance term.

(iv) In the final equation for \( y_t \), the degree of the AR polynomial will be greater than the order of the MA disturbance term given that \( r_{12} + r_{33} > q_{33} \) and \( r_{12} > 0 \). They will be equal if \( r_{12} = 0 \) and \( r_{33} \geq q_{33} \) or if \( r_{12} + r_{33} = q_{33} \). The latter will be greater if \( r_{12} + r_{33} < q_{33} \).

In what follows, post-World War II quarterly data for the U.S., 1947–1972, are employed to analyze the final equations for \( c_t, y_t \), and \( x_t \), and to check some of the implications mentioned above.

From (3.8), the dynamic structural equations of the dynamized Haavelmo model are:

\[
\begin{pmatrix}
1 & h_{12} & 0 \\
h_{21} & 1 & h_{23}
\end{pmatrix}
\begin{pmatrix}
c_t \\
y_t \\
x_t
\end{pmatrix} =
\begin{pmatrix}
\theta_1 \\
\theta_2
\end{pmatrix} +
\begin{pmatrix}
f_{11} & f_{12} \\
f_{21} & f_{22}
\end{pmatrix}
\begin{pmatrix}
e_{1t} \\
e_{2t}
\end{pmatrix},
\]

(3.12a)
or
\[
\begin{pmatrix}
1 & h_{12} \\
1 & 0
\end{pmatrix}
\begin{pmatrix}
c_t \\
y_t
\end{pmatrix}
= \begin{pmatrix}
\theta_1 \\
\theta_2
\end{pmatrix}
+ \begin{pmatrix}
0 \\
-h_{23}
\end{pmatrix}x_t + \begin{pmatrix}
f_{11} & f_{12} \\
f_{21} & f_{22}
\end{pmatrix}
\begin{pmatrix}
e_{1t} \\
e_{2t}
\end{pmatrix}.
\]  
(3.12b)

From (3.12b), the transfer equations, the analogue of (2.15) are:

\[
\begin{pmatrix}
1 & h_{12} \\
1 & 0
\end{pmatrix}
\begin{pmatrix}
c_t \\
y_t
\end{pmatrix}
= \begin{pmatrix}
\theta_1'' \\
\theta_2''
\end{pmatrix}
+ \begin{pmatrix}
1 & -h_{12} \\
-h_{21} & 1
\end{pmatrix}
\begin{pmatrix}
x_t \\
e_{1t}
\end{pmatrix}
+ \begin{pmatrix}
f_{11} & f_{12} \\
f_{21} & f_{22}
\end{pmatrix}
\begin{pmatrix}
e_{1t} \\
e_{2t}
\end{pmatrix},
\]  
(3.13)

or

\[
(1-h_{12}h_{21})c_t = \theta_1' + h_{12}h_{23}x_t + (f_{11} - f_{21}h_{12})e_{1t}
+ (f_{12} - f_{22}h_{12})e_{2t},
\]  
(3.14)

and

\[
(1-h_{12}h_{21})y_t = \theta_2' - h_{23}x_t + (f_{21} - f_{11}h_{21})e_{1t}
+ (f_{22} - f_{12}h_{21})e_{2t},
\]  
(3.15)

where \(\theta_1''\) and \(\theta_2''\) are constant parameters that are linear functions of \(\theta_1\) and \(\theta_2\).

The following properties of the transfer equations, (3.14) and (3.15) are of interest:

(a) The AR parts of the two transfer equations are identical. Since \(h_{12}\) is of degree \(r_{12}\) and \(h_{21}\) of degree \(r_{21}\), the order of the autoregression in each equation is \(r_{12} + r_{21}\).

(b) In (3.14) the degree of the operator \(h_{12}h_{23}\) hitting \(x_t\) is \(r_{12} + r_{23} = r_{12} + r_{21}\), the same as that for the autoregressive part of the equation, \(1-h_{12}h_{21}\).

(c) In (3.15), the degree of the lag operator, \(-h_{23}\), applied to \(x_t\), is \(r_{23} = r_{21}\), which is less than or equal to the degree of \(1-h_{12}h_{21}\), the AR polynomial.

(d) The lag operator acting on \(x_t\) in the equation for \(c_t\), \(h_{12}h_{33}\), is a multiple of that acting on \(x_t\) in the equation for \(y_t\), and thus the former has degree larger than or equal to that of the latter.

(e) If the structural disturbance terms are serially uncorrelated, i.e. \(f_{ij}\) has degree zero in \(L\) for \(i,j = 1,2\), the orders of the MA error terms in (3.14) and (3.15) are \(r_{12} \geq 0\) and \(r_{21} \geq 0\), respectively. Thus for both equations, the order of the MA error process is less than or equal to the order of the AR part of the equation.

By use of appropriate statistical techniques and data, the transfer equations in (3.14)–(3.15) can be analyzed to determine the degrees of lag polynomials and to estimate parameter values. With these results in hand, it is possible to check the points (a)–(e) relating to the transfer equations associated with Haavelmo's dynamic model. Below the results of such calculations are reported.
4. Empirical analyses of final equations (3.9)–(3.11)

4.1. Analyses utilizing Box–Jenkins (BJ) techniques

In this subsection, we report the results of applying BJ identification and estimation procedures to the final equations of the dynamized Haavelmo model. Box and Jenkins (1970, p. 175) provide the following relations between the autocorrelation and partial autocorrelation functions associated with stationary stochastic processes for a single random variable:

1. For a purely AR process of order \( p \), the autocorrelation function tails off and the partial autocorrelation function has a cut-off after lag \( p \). 
2. For a purely MA process of order \( q \), the autocorrelation function has a cut-off after lag \( q \) and the partial autocorrelation function tails off.
3. For a mixed ARMA process, with the order of the AR being \( p \) and that of the MA being \( q \), the autocorrelation function is a mixture of exponential and damped sine waves after the first \( q-p \) lags and the partial autocorrelation function is dominated by a mixture of exponentials and damped sine waves after the first \( p-q \) lags.

Box and Jenkins suggest differencing a series until it is stationary and then computing estimates of the autocorrelation and partial autocorrelation functions. Using (1)–(3), it may be possible to determine or identify the nature of the process for the differenced series as well as values of \( p \) and \( q \). Once the process or model and \( p \) and \( q \) have been determined, the model’s parameters can be estimated, usually by use of a non-linear estimation procedure.

Plots of the data for the variables of Haavelmo’s model, \( c_t, y_t, \) and \( x_t \), are shown in fig. 1. From this figure, it is seen that the variables apparently have trends and thus are non-stationary. First or second differencing of the variables may induce stationarity. For the reader’s benefit, plots of the first differences of the variables are presented in fig. 2. It is clear from the plots of the first differences that they are less subject to trend than are the levels of the variables. However, a slight trend in the magnitudes of the first differences may be present if the levels are subject to a relatively constant proportionate rate of growth. For this reason, we also performed analyses based on second differences.

In fig. 3, we present the estimated autocorrelation function for the series \( c_t - c_{t-1} \), the first difference of consumption. Also indicated in fig. 3, is a \( \pm 2\sigma \) confidence band for the autocorrelations where \( \sigma \) is a large sample standard

---

\(^8\)See Box and Jenkins (1970, pp. 64–65) for definition of this function.

\(^9\)Autocorrelation functions have been formerly used in econometrics, see e.g. Wold (1953).

\(^10\)The variables have been defined above. The data are seasonally adjusted quarterly, price-deflated, per capita aggregates, expressed in dollars at an annual rate, for the U.S. economy, 1947I–1977II, obtained from official sources cited in the appendix.

\(^11\)The computer program employed was developed by C.R. Nelson and S. Beveridge, Graduate School of Business, University of Chicago.
Fig. 1

Plots of Data for $c$, $\gamma$, and $x$

Fig. 2

Plots of Data for $c$, $\gamma$, and $x$

a. $c - c_{t-1}$

b. $\gamma - \gamma_{t-1}$

c. $x - x_{t-1}$
error for the sample autocorrelations. It is seen that all estimated autocorrelations lie within the band except for that of lag 2. This suggests that the underlying process is not purely AR. If the autocorrelation estimate for lag 2 is regarded as a cut-off, the results suggest that a second order MA process may be generating the first differences of $c_t$. The estimated partial autocorrelation function, also shown in fig. 3, does not appear to contradict this possibility. Estimation of a second order MA model for the first differences of consumption, led to the following results using the BJ non-linear algorithm:

$$c_t - c_{t-1} = e_t + 0.021 e_{t-1} + 0.278 e_{t-2} + 10.73 \quad s^2 = 530, \quad (4.1)$$

where $s^2$ is the residual sum of squares (RSS) divided by the number of degrees of freedom and the figures in parentheses are large sample standard errors.

For income, $y_t$, a plot of the first differences is given in fig. 2. From the plot of the estimated autocorrelations for the first differences in fig. 4, it appears that none of the autocorrelations are significantly different from zero, a finding that leads to the presumption that the underlying model is not AR. Estimates of the partial autocorrelations for lag 4 and lag 10 lie close to the limits of the $\pm 2\hat{\sigma}$ band – see fig. 4. Other partial autocorrelations appear not to differ significantly from zero. If all autocorrelations and partial autocorrelations are deemed not significantly different from zero, then the conclusion would be that the first differences of income are generated by a random walk model which was estimated with the following results:

$$y_t - y_{t-1} = e_t + 10.03 \quad s^2 = 842. \quad (4.2)$$

For the first differences of investment, $x_t - x_{t-1}$ – see plot in fig. 2 – the estimated autocorrelation and partial autocorrelation functions are given in fig. 5. The autocorrelations alternate in sign and show some significant values for lags less than or equal to 5 which suggests an AR model. The partial autocorrelation function has a cut-off at lag 4 supporting the presumption that the model is AR and indicating a 4th order AR scheme. Also, the partial autocorrelation function for the second differences has a cut-off at lag 3 while the autocorrelations alternate in sign for lags less than 11, findings which support those derived from analysis of first differences. In view of these findings, a 4th order AR model has been fitted with the data:

$$(1 + 0.263L - 0.0456L^2 + 0.0148L^3 + 0.376L^4)(x_t - x_{t-1}) = e_t + 7.738 \quad s^2 = 939. \quad (4.3)$$

$12\hat{\sigma}^2$ is an estimate of the following approximate variance of $r_k$, the $k$th sample serial correlation, given in Bartlett (1946). With $\rho_v = 0$ for $v > q$, var ($r_k$) $\equiv (1 + 2\sum_{v=1}^{q} \rho_v^2)/T$, for $k > q$. The $\pm 2\hat{\sigma}$ bounds for $r_k$, $k = 1, 2, \ldots, 12$, are calculated under the assumption, $\rho_v = 0, v > 0$. For $k > 12$, they are calculated assuming $\rho_v = 0, v > 12$.  

Fig. 3
Estimated Autocorrelation Function
for $\epsilon_t - \epsilon_{t-1}$

Estimated Partial Autocorrelation Function
for $\epsilon_t - \epsilon_{t-1}$

Fig. 4
Estimated Autocorrelation Function
for $y_t - y_{t-1}$

Estimated Partial Autocorrelation Function
for $y_t - y_{t-1}$

Fig. 5
Estimated Autocorrelation Function
for $x_t - x_{t-1}$

Estimated Partial Autocorrelation Function
for $x_t - x_{t-1}$

Vertical axes of figs. 3–5 on the right: $\hat{\phi}_{kk}$, on the left: $r_t$. 
In contrast to the processes for the first differences of \( c_t \) and \( y_t \) in (4.1) and (4.2), that for the first differences of investment, \( x_t \), in (4.3) has an AR part. Thus the requirement of the structural form that all endogenous variables have identical AR parts of order equal to or greater than that for \( x_t \) — see (3.9)–(3.10) above — is not satisfied given the results in (4.1)–(4.3). Using the notation of table 2 with \( h_{ij} \) of degree \( r_{ij} \) regarded as an element of \( H(L)/(1-L) \), the degree of the AR polynomial in (4.31) is \( r_{33} = 4 \) while that of the error process is \( q_{33} = 0 \).

In the case where no cancelling occurs in (4.1)–(4.2), it is clear that the conditions (3.9) and (3.10) of table 2 can not be met. Even if \( h_{23} \) in (3.8) satisfies \( h_{23} = 0 \) so that \( c_t \) and \( y_t \) are generated independently of \( x_t \), the conditions on the final equations are not met by the results for the final equations in (4.1)–(4.3). Thus while (4.1)–(4.3) appear to be consistent with the information in the data, they are not compatible with the dynamicized Haavelmo model specified in sect. 3, eqs. (3.2a)–(3.2c).

At this point, the following are considerations that deserve attention:

1. Although the fits of the models in (4.1)–(4.3) are fairly good, it may be that schemes somewhat more complicated than (4.1)–(4.3) are equally well or better supported by the information in the data and are compatible with the implications of the Haavelmo model. This possibility is explored below.

2. To compare and test alternative final equations for each variable, it would be desirable to have inference methods that are less 'judgmental' and more systematically formal than are the BJ methods. In the next subsection, we indicate how likelihood ratios and posterior odds ratios can be used for discriminating among alternative final equation models.

3. It must be recognized that there are some limitations on the class of AR models that can be transformed to a stationary process through differencing. That is, only those AR models whose roots lie on the boundary or inside the unit circle can be transformed to stationary models by differencing. Other transformations, say logarithmic, have to be used for models with roots outside the unit circle.

4. Differencing series may amplify the effects of measurement errors present in the original data and seriously affect estimates of the autocorrelation and partial autocorrelation functions. Of course, this problem arises not only in the BJ approach but also in any analysis of ARMA processes, particularly those of high order.

### 4.2. Analyses of final equations utilizing likelihood ratios and posterior odds

The purpose of this section is to provide additional procedures for identifying or determining the forms of final equations. These procedures involve use of

---

13 If \( h_{23} = 0 \), then (4.1)–(4.2) imply \( r_{12} + r_{21} = 1; r_{12} + g_{22}, q_{11}, q_{22}, r_{12} + g_{22} \leq 2 \) (with at least one equality); and \( r_{21} + q_{11}, q_{21}, r_{21} + q_{12}, q_{22} \leq 0 \). These conditions imply \( q_{11} = q_{12} = q_{21} = q_{22} = r_{21} = 0, r_{21} = 1, \) and \( r_{21} = 2 \) which cannot hold simultaneously.
likelihood ratios and Bayesian posterior odds. After showing how to obtain likelihood ratios and posterior odds, some of the results are applied in the analysis of Haavelmo's model.

Consider the following ARMA model for a single random variable $z_i$,

$$\phi(L)z_t = \theta(L)e_t, \quad t = 1, 2, \ldots, T, \quad (4.4)$$

where $\phi(L)$ and $\theta(L)$ are polynomials in $L$ of degree $p$ and $q$, respectively. Assume that the $e_t$'s are normally and independently distributed, each with zero mean and common variance, $\sigma^2$. Let $u_t \equiv \theta(L)e_t$. Then given the 'starting values' for $e_t$ and $z_t$, $e_0$ and $z_0$, the vector $u' = (u_1, u_2, \ldots, u_T)$ has a $T$-dimensional multivariate normal distribution with zero mean vector and covariance matrix $\Sigma$, that is:

$$p(u' | \phi, \theta, \sigma^2, z_0, e_0) = (2\Pi)^{-T/2} \exp \left\| \frac{1}{2} u' \Sigma^{-1} u \right\|, \quad (4.5)$$

where $\phi' = (\phi_1, \phi_2, \ldots, \phi_p)$ and $\theta' = (\theta_1, \theta_2, \ldots, \theta_q)$. The matrix $\Sigma$ is a $T \times T$ positive definite symmetric matrix with elements given by:

$$\begin{align*}
\sigma_{i, t-k} &= \sigma^2 (1 + \sum_{i=1}^{q} \theta_{i}^2), & \text{for } k = 0, \\
\sigma_{i, t-k} &= \sigma^2 (-\theta_k + \sum_{i=k+1}^{q} \theta_{i-k} \theta_{i}), & \text{for } 0 < k \leq q, \\
\sigma_{i, t-k} &= 0, & \text{for } k > q.
\end{align*} \quad (4.6)$$

Also, the joint probability density function (pdf) for the $e_t$'s, given by

$$e_t = z_t - \phi_1 z_{t-1} - \ldots - \phi_p z_{t-p} + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}, \quad (4.7)$$

$$t = 1, 2, \ldots, T,$$

is:

$$p(e | \phi, \theta, \sigma^2, e_0, z_0) = (2\Pi \sigma^2)^{-T/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{T} e_i^2 \right\}. \quad (4.8)$$

Since the Jacobian of the transformation from the $e_t$'s to the $z_t$'s is equal to one, the joint pdf for the $z_t$'s, the likelihood function, is:

$$p(z | \phi, \theta, \sigma^2, e_0, z_0) = (2\Pi \sigma^2)^{-T/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{T} (z_t - \phi_1 z_{t-1} \right.\right.$$  

$$- \ldots - \phi_p z_{t-p} + \theta_1 e_{t-1} + \theta_2 e_{t-2} \left. \right) \right.\right.$$  

$$+ \ldots + \theta_q e_{t-q})^2 \right\}. \quad (4.9)$$
In this context, (4.9) is convenient since Marquardt’s non-linear computational algorithm can be applied to obtain maximum likelihood (ML) estimates.

If we have an alternative ARMA model,

\[ \phi_a(L)z_t = \theta_a(L)e_{at}, \quad t = 1, 2, \ldots, T, \]  

(4.10)

where \( \phi_a(L) \) is of degree \( p_a \), \( \theta_a(L) \) of degree \( q_a \) and the error process \( e_{at} \) is \( NID(0, \sigma_a^2) \), then the likelihood ratio, \( \lambda \), for (4.4) and (4.10) is

\[ \lambda = \left( \max_{\phi, \theta, \sigma} l(\phi, \theta, \sigma|z) \right) / \left( \max_{\phi_a, \theta_a, \sigma_a} l(\phi_a, \theta_a, \sigma_a|z) \right), \]  

(4.11)

where \( l(\phi, \theta, \sigma|z) \) denotes (4.9) viewed as a function of its parameters and similarly for \( l(\phi_a, \theta_a, \sigma_a|z) \). The ratio of maximized likelihood functions in (4.11) reduces to:

\[ \lambda = \left( \frac{\hat{\sigma}_a^2}{\hat{\sigma}^2} \right)^{(2)/2}, \]  

(4.12)

where

\[ \hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} (z_t - \hat{\phi}_1z_{t-1} - \ldots - \hat{\phi}_pz_{t-p})^2 + \hat{\theta}_1\hat{\varepsilon}_{t-1} + \ldots + \hat{\theta}_q\hat{\varepsilon}_{t-q})^2 \]  

(4.13a)

and

\[ \hat{\sigma}_a^2 = \frac{1}{T} \sum_{t=1}^{T} (z_t - \hat{\phi}_{1a}z_{t-1} - \ldots - \hat{\phi}_{pa}z_{t-pa})^2 + \hat{\theta}_{1a}\hat{\varepsilon}_{a,t-1} + \ldots + \hat{\theta}_{qa}\hat{\varepsilon}_{a,t-qa})^2 \]  

(4.13b)

are the ML estimates for \( \sigma^2 \) and \( \sigma_a^2 \).

If model (4.10) is nested in model (4.4), i.e. \( p_a \leq p \) and/or \( q_a \leq q \), with at least one strict inequality, and under the assumption that (4.10) is the true model, \( 2\ln \lambda \) is approximately distributed as \( \chi^2_r \) with \( r \) being the number of restrictions imposed on (4.4) to obtain (4.10); that is, \( r = p + q - (p_a + q_a) \) – see Silvey (1970, pp. 112-113). In choosing a significance level for this test, it is very important, as usual, to consider errors of the first and second kind. Rejecting the nested model when it is ‘true’ appears to us to be a less serious error than failing to reject it when the broader model is ‘true’. That is, using the restricted model when the restrictions are not ‘true’ may lead to serious errors. Use of the broader model, when the restricted model is ‘true’, involves carrying along some extra parameters which may not be as serious a problem as giving these parameters incorrect values. This argues against using extremely low significance levels, e.g. \( \alpha = 0.01 \) or \( \alpha = 0.001 \). Also, these considerations rationalize somewhat the usual practice of some degree of over-fitting when the model form is somewhat uncertain. More systematic analysis and study of this problem would be desirable.
In order to compare (4.4) and (4.10) in a Bayesian context, we have to specify a prior distribution on the parameter space. In the problem of comparing nested models, this prior distribution has a mixed form with weights whose ratio is the prior odds on alternative models — see e.g. Jeffreys (1961, p. 250), Zellner (1971, p. 297ff.) and Palm (1972). Formally, the posterior odds ratio relating to (4.4) and (4.10) is given by:

\[ K_{1a} = \frac{\prod \int p(\phi, \theta, \sigma)p(\phi|\phi_a, \theta_a, \sigma_a) d\phi d\theta d\sigma}{\prod_a \int p(\phi_a, \theta_a, \sigma_a)p(\phi|\phi_a, \theta_a, \sigma_a) d\phi_a d\theta_a d\sigma_a}, \quad (4.14) \]

where \( K_{1a} \) is the posterior odds ratio, \( \Pi/\Pi_a \) is the prior odds ratio, and \( p(\phi, \theta, \sigma) \) and \( p(\phi_a, \theta_a, \sigma_a) \) are the prior pdf’s for the parameters. Before (4.14) can be made operational, it is necessary to formulate the prior pdf’s and to evaluate the integrals, either exactly or approximately.\(^4\)

We now compute likelihood ratios to compare alternative formulations of the final equations of Haavelmo’s model. The information in table 2 and empirical results in the literature on quarterly consumption relations suggest higher order AR and MA schemes than those fitted in sect. 4.1. For example, a 4th order AR model for the 2nd differences of consumption with 3rd order MA error terms is a scheme tentatively suggested by considerations presented in sect. 4.1. This scheme has been fitted with both the consumption and income data with results shown for \( c_t \) and \( y_t \) in tables 3 and 4 and those for \( x_t \) in table 5 with figures in parentheses being large sample standard errors. Also shown in table 3 are the results for the simple schemes of sect. 4.1 and results for several other specifications. It should be noted that use of the broader schemes for \( c_t \) and \( y_t \) results in decreases in the value of the residual sum of squares divided by degrees of freedom of about 8 to 12 per cent. However, it must be noted that the large sample standard errors associated with the point estimates are rather large in number of instances.

To put the comparison of alternative schemes on a more formal basis, likelihood ratios have been computed and are reported in table 6. Using these ratios as a basis for large sample \( \chi^2 \) tests, it is found that it is possible to reject the simpler versions at reasonable significance levels. The results of the tests indicate that it is reasonable to retain the model (5, 1, 4) for consumption and income and

\(^4\)Note however, as pointed out by Lindley (1961), the likelihood functions in the numerator and denominator of (4.14) can be expanded about ML estimates. If just the first terms of these expansions are retained, namely \( l(\hat{\phi}, \hat{\theta}, \hat{\sigma}|z) \) and \( l(\hat{\phi_a}, \hat{\theta_a}, \hat{\sigma_a}|z) \), and if the prior pdf’s are proper, (4.14) is approximated by:

\[ K_{1a} \approx \frac{\Pi/\Pi_a l(\hat{\phi}, \hat{\theta}, \hat{\sigma}|z)/l(\hat{\phi_a}, \hat{\theta_a}, \hat{\sigma_a}|z)}, \]

i.e. a prior odds ratio, \( \Pi/\Pi_a \), times the usual likelihood ratio. As Lindley points out, additional terms in the expansions can be retained and the resulting expression will involve some prior moments of parameters. Thus on assigning a value to \( \Pi/\Pi_a \), the prior odds ratio, the usual likelihood ratio is transformed into an approximate posterior odds ratio for whatever non-dogmatic, proper prior pdf’s employed.
# Table 3

Estimated final equations for consumption.

<table>
<thead>
<tr>
<th>Model ((p, d, q))</th>
<th>RSS, residual sum of squares</th>
<th>DF</th>
<th>RSS/DF</th>
<th>Estimates of the AR part</th>
<th>Estimates of the MA part</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AR1</td>
<td>AR2</td>
<td>AR3</td>
</tr>
<tr>
<td>1) ((0, 1, 2))</td>
<td>52,012</td>
<td>98</td>
<td>530</td>
<td>-0.021</td>
<td>-0.278</td>
<td>10.727</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.101)</td>
<td>(0.101)</td>
<td></td>
</tr>
<tr>
<td>2) ((4, 2, 3))</td>
<td>45,807</td>
<td>92</td>
<td>497</td>
<td>0.727</td>
<td>0.199</td>
<td>-0.077</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.365)</td>
<td>(0.255)</td>
<td>(0.178)</td>
</tr>
<tr>
<td>3) ((5, 2, 2))</td>
<td>48,411</td>
<td>92</td>
<td>525</td>
<td>-0.651</td>
<td>0.226</td>
<td>-0.0567</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.458)</td>
<td>(0.142)</td>
<td>(0.160)</td>
</tr>
<tr>
<td>4) ((5, 1, 4))</td>
<td>44,634</td>
<td>91</td>
<td>490</td>
<td>0.587</td>
<td>0.514</td>
<td>-0.166</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.101)</td>
<td>(0.190)</td>
<td>(0.246)</td>
</tr>
</tbody>
</table>

\(\(p, d, q\)\) denotes an ARMA model for the \(d\)th differences of a variable that has AR polynomial of degree \(p\) and MA polynomial of degree \(q\).
Table 4
Estimated final equations for income.

<table>
<thead>
<tr>
<th>Model (p, d, q)</th>
<th>RSS, residual sum of squares</th>
<th>DF</th>
<th>RSS/DF</th>
<th>Estimates of the AR part</th>
<th>Estimates of the MA part</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AR1</td>
<td>AR2</td>
<td></td>
</tr>
<tr>
<td>1) (0, 1, 0)</td>
<td>85,042</td>
<td>101</td>
<td>842</td>
<td>0.0803</td>
<td>-0.0156</td>
<td>-0.307</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.353)</td>
<td>(0.112)</td>
<td>(0.110)</td>
</tr>
<tr>
<td>2) (0, 1, 4) + restrictions</td>
<td>79,362</td>
<td>99</td>
<td>797</td>
<td>0.0861</td>
<td>-0.0156</td>
<td>-0.302</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.118)</td>
<td>(0.112)</td>
<td>(0.372)</td>
</tr>
<tr>
<td>3) (4, 1, 1)</td>
<td>78,546</td>
<td>95</td>
<td>823</td>
<td>-0.429</td>
<td>-0.063</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.418)</td>
<td>(0.153)</td>
<td>(0.372)</td>
</tr>
<tr>
<td>4) (4, 2, 3)</td>
<td>78,827</td>
<td>92</td>
<td>821</td>
<td>-0.154</td>
<td>0.222</td>
<td>0.440</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.532)</td>
<td>(0.120)</td>
<td>(0.423)</td>
</tr>
<tr>
<td>5) (4, 1, 4)</td>
<td>65,428</td>
<td>92</td>
<td>703</td>
<td>-0.390</td>
<td>-0.437</td>
<td>-0.295</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.244)</td>
<td>(0.168)</td>
<td>(0.249)</td>
</tr>
<tr>
<td>6) (5, 1, 4)</td>
<td>65,351</td>
<td>91</td>
<td>705</td>
<td>-0.552</td>
<td>-0.390</td>
<td>-0.733</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.380)</td>
<td>(0.244)</td>
<td>(0.215)</td>
</tr>
</tbody>
</table>

*(p, d, q) denotes an ARMA model for the dth differences of a variable that has AR polynomial of degree p and MA polynomial of degree q.*
### Table 5
Estimated final equations for investment.

| Model \((p, d, q)\)* | RSS, residual sum of squares | DF | RSS/DF | AR1 \(\hat{\theta}_1\) | AR2 \(\hat{\theta}_2\) | AR3 \(\hat{\theta}_3\) | AR4 \(\hat{\theta}_4\) | AR5 \(\hat{\theta}_5\) | MA1 \(\hat{\theta}_{1,1}\) | MA2 \(\hat{\theta}_{2,1}\) | MA3 \(\hat{\theta}_{3,1}\) | MA4 \(\hat{\theta}_{4,1}\) | Constant \(\hat{\theta}_0\) |
|-----------------|----------------------------|-----|--------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 1) \((4, 1, 0)\) | 90,519                     | 96  | 939    | -0.263          | 0.0456          | -0.0148         | -0.376          |                  |                  |                  |                  |                  |                  |                  | 7.738 (3.265)   |
|                 |                            |     |        | (0.0942)       | (0.0976)        | (0.0970)        | (0.0933)        |                  |                  |                  |                  |                  |                  |                  |
| 2) \((4, 2, 3)\) | 91,124                     | 92  | 978    | -0.653          | -0.183          | -0.0334         | -0.325          |                  |                  |                  |                  |                  |                  |                  | -0.127 (0.232)  |
|                 |                            |     |        | (0.284)         | (0.294)         | (0.127)         | (0.127)         |                  |                  |                  |                  |                  |                  |                  |
| 3) \((5, 2, 1)\) | 87,428                     | 93  | 929    | -0.247          | 0.0395          | -0.0245         | -0.337          | 0.106           |                  |                  |                  |                  |                  |                  | 1.014 (0.102)   |
|                 |                            |     |        | (0.0924)       | (0.0991)        | (0.0980)        | (0.0991)        | (0.096)         |                  |                  |                  |                  |                  |                  |
| 4) \((5, 2, 2)\) | 90,391                     | 92  | 970    | -0.659          | -0.076          | -0.0076         | -0.364          | -0.065          | 0.552           | 0.425           |                  |                  |                  |                  | -1.04 (0.187)   |
|                 |                            |     |        | (0.631)         | (0.192)         | (0.127)         | (0.118)         | (0.276)         | (0.609)         | (0.630)         |                  |                  |                  |                  |

*\((p, d, q)\) denotes an ARMA model for the \(d\)th differences of a variable that has AR polynomial of degree \(p\) and MA polynomial of degree \(q\).*
(4, 1, 0) for investment. Given that these models are tentatively accepted, it is the case that the AR and MA polynomials for the consumption and income processes have identical degrees. However, the point estimates of the AR parameters of consumption and income processes are not very similar, a finding that must be tempered by the fact that standard errors associated with coefficient estimates are rather large, particularly for the AR parameters of the income process. It would be very desirable to develop joint estimation techniques for the two processes.

Table 6
Results of large sample likelihood ratio tests applied to final equations of Haavelmo’s model.

| Models compared | $\lambda = \frac{L(X|H_1)}{L(X|H_0)}$, $2 \ln \lambda$ | Critical points for $\chi^2_\nu$ |
|-----------------|-----------------------------------|-----------------|
|                 | $\nu = 0.05$ | $\nu = 0.10$ | $\nu = 0.20$ |
| 1. Consumption $c_t$ | | | |
| $H_0: (0, 1, 2)$ vs. $H_1: (4, 2, 3)$ | 573.547 | - | - | - |
| $H_0: (0, 1, 2)$ vs. $H_1: (5, 1, 4)$ | 2098.29 | 15.230 | 7 | 14.07 | 12.02 | 9.80 |
| $H_0: (5, 2, 2)$ vs. $H_1: (4, 2, 3)$ | 15.871 | - | - | - | - |
| $H_0: (4, 2, 3)$ vs. $H_1: (5, 1, 4)$ | 3.654 | 2.592 | 2 | 5.99 | 4.61 | 3.22 |
| 2. Income $y_t$ | | | |
| $H_0: (0, 1, 0)$ vs. $H_1: (0, 1, 4)$ | 31.697 | 6.912 | 1 | 3.84 | 2.71 | 1.64 |
| $H_0: (0, 1, 0)$ vs. $H_1: (4, 1, 4)$ | 4937.0 x 10^2 | 26.219 | 8 | 15.51 | 13.36 | 11.03 |
| $H_0: (0, 1, 0)$ vs. $H_1: (5, 1, 4)$ | 5236.0 x 10^2 | 26.337 | 9 | 16.92 | 14.68 | 12.24 |
| $H_0: (0, 1, 0)$ vs. $H_1: (4, 2, 3)$ | 44.453 | - | - | - | - |
| $H_0: (4, 2, 3)$ vs. $H_1: (5, 1, 4)$ | 117.8 x 10^2 | 18.748 | 2 | 5.99 | 4.61 | 3.22 |
| 3. Investment $x_t$ | | | |
| $H_0: (4, 1, 0)$ vs. $H_1: (5, 2, 1)$ | 5.678 | - | - | - | - | - |

$H: (p, d, q)$ denotes an ARMA model for the $d$th difference of a variable that has AR polynomial of degree $p$ and MA error polynomial of degree $q$.

$^b$These are non-nested hypotheses.

$^c$Here there are 3 restrictions on the parameters of the MA error process.

Equations in order to increase the precision of estimation and joint test procedures for testing the hypothesis that the AR parameters are the same for the two processes.

What are the implications of retaining (5, 1, 4) models for $c_t$ and $y_t$ and a (4, 1, 0) model for $x_t$? As noted above, the empirical finding that first differencing appears adequate to induce stationarity for all three variables suggests that the model can be expressed in first difference form. That is, we rewrite (3.5) as follows:

$$\bar{H}(L)(1-L)c_t = \theta + F(L)e_t,$$

where $\bar{H}(L)$ has elements that are the elements of $H(L)$ divided by $1-L$. With
the polynomials $h_{ij}(L)$ of degree $r_{ij}$ considered elements of $\tilde{H}(L)$ rather than $H(L)$ and if no cancelling occurs in (4.15), then under the restrictions imposed on the Haavelmo model in the preceding section (see table 2), we have $r_{33} = 4$; $q_{33} = 0$; $r_{12} + r_{21} + r_{33} = 5$;

$$r_{33} + q_{11}, r_{33} + r_{12} + q_{21}, r_{33} + q_{12}, r_{33} + r_{12} + q_{22},$$

$$r_{12} + r_{23} + q_{32} \leq 4$$

and

$$r_{33} + q_{21}, r_{33} + r_{21} + q_{11}, r_{33} + q_{22}, r_{33} + r_{21} + q_{12}, r_{23} + q_{33} \leq 4,$$

with at least one equality holding in both cases. These restrictions imply $r_{12} + r_{21} = 1$, all $q_{ij} = 0$ and $r_{12} = r_{21} = 0$, conditions that cannot hold simultaneously. Also, if we retain a $(5, 2, 1)$ model for investment, we end up with a contradiction.

If we make the assumption that the joint process for $\Delta c_t$ and $\Delta y_t$ is independent of $\Delta x_t$, i.e. $h_{23}(L) \equiv 0$ in (3.8), an assumption that may appeal to some Quantity of Money theorists but not to most Keynesians, the degrees of the polynomials reported in table 2 are reduced by $r_{33}$ and we have the following restrictions on the degrees of the AR polynomials in the processes for $\Delta c_t$ and $\Delta y_t$: $r_{12} + r_{21} = 5$; $q_{11}, q_{12}, r_{12} + q_{21}, r_{12} + q_{22} \leq 4$; and $q_{21}, q_{22}, r_{21} + q_{11}, r_{21} + q_{12} \leq 4$, with at least one equality holding in both cases. With further assumptions, e.g. $r_{12} = 2$ and $r_{21} = 3$, it is possible to determine compatible values for the degrees of the structural equations' lag polynomials and error term polynomials. However, this compatibility is attained only with the controversial assumption that the joint process for $\Delta c_t$ and $\Delta y_t$ is independent of the process for $\Delta x_t$, Haavelmo's investment variable. A major implication of this last assumption is that the analysis of the transfer functions should reveal no dependence of either $\Delta c_t$ or $\Delta y_t$ on $\Delta x_t$, a point that is checked in the next section where we analyze the transfer equations (3.14)-(3.15).

An alternative way to achieve compatibility of the results of the final equation analyses with structural assumptions, is to assume that $h_{23}(L) \equiv h_{33}(L)$. This assumption implies that the investment variable influences $\Delta c_t$ and $\Delta y_t$ only through its disturbance term. With this assumption, $h_{33}$ cancels in eqs. (3.9) and (3.10) and the empirical findings combined with the results in table 2, imply that $r_{33} = 4, r_{23} = r_{33} = 4$, by assumption, $q_{33} = 0, r_{12} + r_{21} = 5$,

$$q_{11}, q_{12}, r_{12} + q_{21}, r_{12} + q_{22} \leq 4$$

and

$$q_{21}, q_{22}, r_{21} + q_{11}, r_{21} + q_{12} \leq 4,$$

with at least one equality in each case. Further, the autoregressive parts of the

15In terms of (3.8), this assumption implies that $h_{23}(L) \equiv 0$. With this assumption, the structural equations are, from (3.3): $\Delta c_t = \alpha(L)\Delta y_t + \beta + u$ and $\Delta y_t = (1 - \mu(L))\Delta c_t - \nu - w$. That is, current and lagged values of $y_t$ affect consumption and current and lagged consumption affect income.
final equations for \( c_t \) and \( y_t \) are identical to the autoregressive parts in their transfer equations. These implications of the assumption, \( h_{23}(L) = h_{33}(L) \), and of final equation findings for the forms of the transfer equations will be checked in the next section.

5. Empirical analyses of transfer equations (3.14)–(3.15)

We now turn to an analysis of the transfer functions, shown in (3.14)–(3.15), associated with the dynamized Haavelmo model. These equations express \( c_t \) and \( y_t \) as functions of their own lagged values, of current and lagged values of \( x_t \), and of current and lagged error terms. The first step in the analysis of the transfer functions is the determination or identification of the degrees of the lag polynomials. In general a transfer function can be written as an infinite moving average process in exogenous variables plus an error term, \( u_t \), with zero mean, that is,

\[
y_t = v(L)x_t + u_t,
\]

where \( v(L) = \sum_{i=0}^{\infty} v_i L^i \). Often this infinite process can be well approximated by a finite distributed lag model of order \( k \), that is \( v(L) = \sum_{i=0}^{k} v_i L^i \). Solving the Yule–Walker equations for a \( k \)th order approximation, we obtain

\[
E_{x_t x_{t-\tau}} = v_0 E_{x_{t-\tau}} + v_1 E_{x_{t-1} x_{t-\tau}} + \ldots + v_k E_{x_{t-k} x_{t-\tau}}, \quad \tau = 0, 1, 2, \ldots, k.
\]

Rough estimates of the \( v_i \)'s can be obtained by replacing the expectations in (5.2) by corresponding sample moments and solving for the \( v_i \)'s. This is equivalent to regressing \( y_t \) on current and lagged values of \( x_t \). Note too that \( v(L) \) can be written as the ratio of two lag polynomials of degrees \( s \) and \( r \), \( \omega_s(L) \) and \( \theta_r(L) \), as follows:

\[
v(L) = \omega_s(L)/\theta_r(L) [L^b],
\]

with \( b \) some non-negative integer. Introduction of \( b \neq 0 \) allows for some ‘dead time’ in the response pattern of \( y_t \) to \( x_t \). Using (5.3) and the preliminary estimates of the \( v_i \)'s, obtained as described above, preliminary estimates of the parameters \( \omega_s(L), \omega_j \)'s, \( j = 0, 1, 2, \ldots, s \), and of \( \theta_r(L) \), \( \theta_i \)'s, \( i = 1, 2, \ldots, r \), can be found. As Box and Jenkins (1970, p. 378) point out, the \( v_j \)'s, the impulse response weights consist of:

1. \( b \) zero values, \( v_0, v_1, \ldots, v_{b-1} \),
2. a further \( s-r+1 \) values, \( v_b, \ldots, v_{b+s-r} \), following no fixed pattern (if \( s < r \), no such values occur), and
3. values \( v_j \), with \( j \geq b+s-r+1 \), following a pattern given by an \( r \)th order difference equation with starting values \( v_{b+s} \ldots v_{b+s-r+1} \).

Properties (1)–(3) can help to determine the values of \( b, s \) and \( r \) from the preliminary estimates of the \( v \)'s. Then the residuals \( \hat{y}_t = y_t - b(L)x_t \) are analyzed to
determine the degrees of the AR and MA parts of the error process using estimated autocorrelation and partial autocorrelation functions. Final estimation of the transfer function so determined can be accomplished in the BJ approach by use of Marquardt's non-linear algorithm.

It is important to observe that the results of final equation analyses can be employed to obtain some information about the degrees of transfer functions' lag polynomials. In fact, if assumptions regarding structural equations' forms are in accord with information in the data, there should be compatibility between the final equations' and transfer equations' forms that we determine from the data.\(^{16}\) That is, final equation analysis led us to (5, 1, 4) processes for \(c_t\) and \(y_t\) and to a (4, 1, 0) process for \(x_t\), namely, \(\phi_{(4)}(L)\Delta x_t = e_t\), or \(\Delta x_t = \phi_{(4)}^{-1}(L)e_t\). If we difference the transfer functions in (3.14) and (3.15) and then substitute \(\Delta x_t = \phi_{(4)}^{-1}(L)e_t\), we obtain the final equations for \(c_t\) and \(y_t\). To obtain compatibility with the empirically determined (5, 1, 4) final equations for \(c_t\) and \(y_t\), the transfer functions must have polynomials hitting \(\Delta x_t\) with degree \(s \leq 3\) in the numerator and degree \(r = 0\) in the denominator. In addition, the ratio of lag polynomials operating on the transfer functions' error terms should have a numerator of degree zero and denominator of degree one.

Under the assumption \(h_{23}(L) \equiv h_{32}(L)\), introduced tentatively in the previous section, the final equation analyses yield the following implications for the transfer functions' lag structures in (3.14)-(3.15):

1. The AR parts of both transfer functions are identical with the AR parts of the final equations and are of degree \(r_{12} + r_{21} = 5\).
2. The polynomial, \(h_{23}\), hitting \(\Delta x_t\) in the transfer function for income is identical with the AR part of the final equation for \(x_t\) and has degree \(r_{23} = r_{33} = 4\).
3. The polynomial operating on \(\Delta x_t\) in the consumption transfer function has degree of at least 4.
4. The order of the moving average error process in each transfer equation is equal to 4.

We shall check points (1)-(4) in the empirical analyses that follow. In this connection, it is the case that there is no assurance that the information in the data will be in accord with compatible findings for the final equations and transfer functions of the Haavelmo model since model specification errors, measurement errors, imperfect seasonal adjustment, etc., can affect analyses to produce incompatible results.

To determine the degrees of the lag polynomials in (3.14)-(3.15) and to get starting values for the \(v_i\)'s, different values for \(k\) were employed in connection with (5.1)-(5.2) that provided preliminary estimates of the \(v_i\)'s. For \(k = 8\) and the first difference of \(c_t\), the \(v_i\)'s with \(i = 0, 1\) and 6 appear to be significantly

\(^{16}\)Here we abstract from the possibility that \(h_{23}(L) = 0\) since in this case transfer functions show no dependence on \(x_t\), a point that is checked below.
different from zero and the behavior of the estimated $v_i$'s is very irregular. The fact that $v_0$ is significantly different from zero implies that $b = 0$. With respect to determining values for $r$ and $s$, the degrees of the polynomials in (5.3), the results are not very precise. The values indicated by our final equation analysis are used and starting values for the $\omega_i$'s are based on the estimates of the $v_i$'s for alternative values of $s$ and $r$. Further, the analysis of the residuals from an 8th order distributed lag model for the first difference of $c_t$ suggests a mixed first order AR and second order MA error process. However, it is thought that this determination of the transfer function's properties is very tentative and thus it was thought worthwhile to proceed to estimate transfer functions in forms suggested by our final equation analyses. Some estimation results for these forms are shown in (5.4)-(5.5): with $s = 3$,

$$
\Delta c_t = (-0.129 + 0.188L + 0.0875L^2 - 0.037L^3)\Delta x_t + e_t/(1 + 0.0208L) + 10.41,
$$

with residual sum of squares (RSS) equal to 41,617, and with $s = 2$,

$$
\Delta c_t = (-0.149 + 0.172L + 0.0830L^2)\Delta x_t + e_t/(1 - 0.0047L) + 10.03,
$$

with RSS = 43,226.

Under the assumption $h_{23}(L) \equiv h_{33}(L)$, an estimate of the transfer function form suggested by the final equation analysis is:

$$
\Delta c_t = \frac{0.0349 + 0.171L + 0.264L^2}{1 + 0.575L - 0.187L^2}\Delta x_t + e_t/1 - 0.550L - 0.559L^2 + 1 - 0.484L - 0.695L^2 + 0.288L^3e_t + 10.55,
$$

with RSS = 30,945.

With respect to the first differences of $y_t$, with $k = 8$, implementation of (5.1)-(5.2) resulted in just $v_0$ being significantly different from zero suggesting that $b = 0$. The estimated $v_i$'s appear to follow a damped wave-like pattern. The difference between the values of $r$ and $s$ is thus thought to be small but this inference is very uncertain. In view of this, the values of $s$ and $r$ implied by the final equation analysis, $s \leq 3$ and $r = 0$ have been employed along with a ratio of polynomials for the error process with numerator of degree 0 and denominator of degree 1. Some estimates reflecting these considerations follow.
For $s = 3$,
\[
\Delta y_t = (0.385 + 0.139L + 0.0938L^2 - 0.124L^3)\Delta x_t + \frac{c_t}{1 + 0.055L} + 10.83,
\]

with RSS = 38,908, and for $s = 2$,
\[
\Delta y_t = (0.355 + 0.121L + 0.117L^2)\Delta x_t + c_t(1 + 0.0077L) + 10.01,
\]

with RSS = 41,547.

Under the assumption that $h_{13}(L) = h_{23}(L)$, the transfer function for income suggested by the final equation analyses has been estimated with the following results:

\[
\Delta y_t = \frac{0.417 + 0.063L + 0.068L^2}{1 - 0.0628L - 0.111L^2 + 0.389L^3}\Delta x_t + 1 + 0.517L + 1 + 0.551L - 0.147L^2c_t + 11.02
\]

with RSS = 36,462.

The estimates reported in (5.4)-(5.9) are in accord with the implications of final equation analyses for the forms of the transfer functions. Further, we see that for (5.4)-(5.5) and (5.7)-(5.8), the AR polynomials for $\Delta c_t$ and $\Delta y_t$ are almost identical, a requirement that the transfer functions must satisfy given that the variables are generated by a joint process with $x_t$ exogenous. Further, from (3.14), (a) the AR part of the transfer function for $\Delta c_t$ should be identical, up to degree 1, to that operating on $\Delta x_t$ in the same equation if the Haavelmo model is adequate, and (b) the lag operator acting on $\Delta x_t$ in the equation for $\Delta c_t$ should be a multiple of that for $\Delta x_t$ in the income equation. The first of these requirements is not satisfied by the results in (5.4)-(5.5) since the polynomials acting on $\Delta c_t$ and $\Delta x_t$ have differing degrees. However, the requirements (a) and (b) are satisfied, as far as the degrees are concerned, for (5.6) and (5.9).  

Last, as mentioned above, one way to have the empirically determined final equations compatible with the dynamized Haavelmo model in (3.8) is to assume $h_{23} \equiv 0$, i.e. that the process for $\Delta x_t$ is independent of the joint process for $x_t$. This suggests that the restriction $h_{23} \equiv h_{23}$, originally imposed, is probably not in accord with the information in the data.
Table 7
Estimated transfer functions for consumption.

<table>
<thead>
<tr>
<th>Model</th>
<th>RSS</th>
<th>DF</th>
<th>RSS/DF</th>
<th>Estimates of the AR and MA parts of $\Delta x_i$</th>
<th>Estimates of the AR and MA parts of the error process</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>41,617</td>
<td>89</td>
<td>467.6</td>
<td>$-0.129 + 0.188L + 0.0875L^2 - 0.037L^3$</td>
<td>$1/(1 + 0.0208L)$</td>
<td>10.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>($0.0689$) ($0.0704$) ($0.070$) ($0.068$)</td>
<td>($0.105$)</td>
<td></td>
</tr>
<tr>
<td>$M_2$</td>
<td>43,286</td>
<td>91</td>
<td>475.0</td>
<td>$-0.149 + 0.172L + 0.0830L^2$</td>
<td>$1/(1 - 0.0047L)$</td>
<td>10.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>($0.0685$) ($0.0709$) ($0.0677$)</td>
<td>($0.105$)</td>
<td></td>
</tr>
<tr>
<td>$M_3$</td>
<td>40,212</td>
<td>90</td>
<td>446.8</td>
<td>$0.370 + 0.126L + 0.0829L^2 - 0.114L^3$</td>
<td>$1 - 0.0376L$</td>
<td>10.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>($0.067$) ($0.068$) ($0.0677$) ($0.066$)</td>
<td>($0.106$)</td>
<td></td>
</tr>
<tr>
<td>$M_4$</td>
<td>37,941</td>
<td>84</td>
<td>451.7</td>
<td>$-0.080 + 0.085L + 0.219L^2$</td>
<td>$1 + 0.495L$</td>
<td>10.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>($0.074$) ($0.108$) ($0.0897$)</td>
<td>($0.530$)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1 + 0.717L - 0.143L^2$ ($0.351$) ($0.349$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_5$</td>
<td>36,504</td>
<td>82</td>
<td>445.2</td>
<td>$-0.111 + 0.183L + 0.0526L^2 + 0.123L^3$</td>
<td>$1 - 0.386L$</td>
<td>10.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>($0.0699$) ($0.0767$) ($0.083$)</td>
<td>($0.334$)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1 + 0.111L + 0.794L^2$ ($0.140$) ($0.114$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_6$</td>
<td>37,645</td>
<td>85</td>
<td>442.3</td>
<td>$-0.129 + 0.050L + 0.187L^2$</td>
<td>$1 + 0.530L$</td>
<td>10.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>($0.072$) ($0.103$) ($0.076$)</td>
<td>($0.219$)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1 + 0.838L + 0.275L^2 + 0.392L^3$ ($0.246$) ($0.331$) ($0.203$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_7$</td>
<td>33,453</td>
<td>83</td>
<td>403.1</td>
<td>$-0.0349 + 0.171L + 0.264L^2$</td>
<td>$1 - 0.550L - 0.559L^2$ ($0.388$) ($0.465$)</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>($0.068$) ($0.106$) ($0.102$)</td>
<td>($0.327$) ($0.431$) ($0.200$)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1 - 0.575L - 0.187L^2$ ($0.269$) ($0.480$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1 - 0.484L - 0.695L^2 + 0.728L^3$ ($0.327$) ($0.431$) ($0.200$)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8
Estimated transfer functions for income.

<table>
<thead>
<tr>
<th>Model</th>
<th>RSS</th>
<th>DF</th>
<th>RSS/DF</th>
<th>Estimates of the AR and MA parts of $\Delta x^i$</th>
<th>Estimates of the AR and MA parts of the error process</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>38,908</td>
<td>89</td>
<td>437.2</td>
<td>$0.385 + 0.139L + 0.0938L^2 - 0.124L^3$</td>
<td>$1/(1 + 0.055L)$</td>
<td>10.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.15)</td>
</tr>
<tr>
<td>$M_2$</td>
<td>41,547</td>
<td>91</td>
<td>456.6</td>
<td>$0.355 + 0.121L + 0.117L^2$</td>
<td>$1/(1 - 0.0077L)$</td>
<td>10.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.301)</td>
</tr>
<tr>
<td>$M_3$</td>
<td>43,149</td>
<td>90</td>
<td>479.4</td>
<td>$-0.145 + 0.173L + 0.075L^2 - 0.0265L^3$</td>
<td>$1 - 0.0127L$</td>
<td>10.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.34)</td>
</tr>
<tr>
<td>$M_4$</td>
<td>43,751</td>
<td>92</td>
<td>475.6</td>
<td>$-0.156 + 0.164L + 0.088L^2$</td>
<td>$1 + 0.0046L$</td>
<td>9.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.33)</td>
</tr>
<tr>
<td>$M_5$</td>
<td>37,775</td>
<td>87</td>
<td>434.2</td>
<td>$0.408 + 0.064L$</td>
<td>$1 - 0.228L$</td>
<td>11.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.63)</td>
</tr>
<tr>
<td>$M_6$</td>
<td>42,819</td>
<td>89</td>
<td>481.1</td>
<td>$0.390 + 0.110L + 0.180L^2$</td>
<td>$1 - 0.181L - 0.172L^2$</td>
<td>10.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.301)</td>
</tr>
<tr>
<td>$M_7$</td>
<td>36,462</td>
<td>85</td>
<td>429.0</td>
<td>$0.417 + 0.063L + 0.068L^2$</td>
<td>$1 + 0.517L$</td>
<td>11.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.396)</td>
</tr>
</tbody>
</table>
\( \Delta c_t \) and \( \Delta y_t \). This implies no dependence of \( \Delta c_t \) and of \( \Delta y_t \) on \( \Delta x_t \) in the transfer equations. The dependence that has been found above might be interpreted as due to specification errors (e.g. \( x_t \) might not be exogenous) or to other complicating factors (e.g. measurement errors, poor seasonal adjustment, etc.). On the other hand, it may be that the alternative assumption \( h_{23} \equiv h_{33} \) is more in accord with the information in the data. Note the substantial reduction in RSS associated with (5.6) and (5.9) relative to the RSS for other models.

Table 9

Results of large sample likelihood ratio tests applied to transfer functions of Haavelmo's model.

| Models compared | \( \lambda = \frac{L(y|H_1)}{L(y|H_0)} \) | 2 ln \( \lambda \) | \( r \) | Critical points for \( \chi^2 \) |
|------------------|-------------------------------------|-----------------|-------|-------------------------------|
|                  |                                     |                 |       | \( \alpha = 0.05 \) | \( \alpha = 0.10 \) | \( \alpha = 0.20 \) |
| 1. Consumption   |                                     |                 |       |                               |                               |                               |
| (1) \( H_0: M_2 \) | \( H_1: M_1 \)                      | 6.66            | 3.79  | 1                             | 3.84                         | 2.71                         | 1.64                         |
| (2) \( H_0: M_1 \) | \( H_1: M_5 \)                      | 702.27          | 13.11 | 5                             | 11.07                        | 9.24                         | 7.29                         |
| (3) \( H_0: M_1 \) | \( H_1: M_4 \)                      | 678.89          | 13.04 | 5                             | 11.07                        | 9.24                         | 7.29                         |
| (4) \( H_0: M_2 \) | \( H_1: M_5 \)                      | 4679.79         | 16.90 | 6                             | 12.59                        | 10.64                        | 8.56                         |
| (5) \( H_0: M_2 \) | \( H_1: M_6 \)                      | 992.41          | 13.80 | 5                             | 11.07                        | 9.24                         | 7.29                         |
| (6) \( H_0: M_3 \) | \( H_1: M_5 \)                      | 126.11          | 9.67  | 5                             | 11.07                        | 9.24                         | 7.29                         |
| (7) \( H_0: M_4 \) | \( H_1: M_5 \)                      | 6893            | 3.86  | 1                             | 11.07                        | 9.24                         | 7.29                         |
| (8) \( H_0: M_3 \) | \( H_1: M_7 \)                      | \( 3.5 \times 10^5 \) | 25.56 | 5                             | 11.07                        | 9.24                         | 7.29                         |
| (9) \( H_0: M_4 \) | \( H_1: M_7 \)                      | \( 5.02 \times 10^2 \) | 12.44 | 1                             | 3.84                         | 2.71                         | 1.64                         |
| 2. Income       |                                     |                 |       |                               |                               |                               |                               |
| (1) \( H_0: M_2 \) | \( H_1: M_1 \)                      | 26.609          | 6.563 | 1                             | 3.84                         | 2.71                         | 1.64                         |
| (2) \( H_0: M_4 \) | \( H_1: M_3 \)                      | 1.999           | 1.385 | 1                             | 3.84                         | 2.71                         | 1.64                         |
| (3) \( H_0: M_4 \) | \( H_1: M_6 \)                      | 2.935           | 2.153 | 3                             | 7.81                         | 6.25                         | 4.64                         |
| (4) \( H_0: M_2 \) | \( H_1: M_7 \)                      | 683.83          | 13.06 | 5                             | 11.07                        | 9.24                         | 7.29                         |
| (5) \( H_0: M_4 \) | \( H_1: M_7 \)                      | 9,065           | 18.22 | 5                             | 11.07                        | 9.24                         | 7.29                         |
| (6) \( H_0: M_4 \) | \( H_1: M_7 \)                      | 3,088           | 16.07 | 2                             | 5.99                         | 4.61                         | 3.22                         |
| (7) \( H_0: M_4 \) | \( H_1: M_7 \)                      | 5.86            | 3.54  | 1                             | 3.84                         | 2.71                         | 1.64                         |

To explore this last point more systematically, some alternative transfer function models, formulated without taking into account results of final equation analyses, have been estimated with results shown in tables 7 and 8. For comparison, results with models implied by the final equation analyses are also presented. A quick look at the residual sum of squares (RSS) indicates that for consumption, alternative model \( M_4 \) yields about a 12 per cent reduction in RSS relative to \( M_1 \) while \( M_7 \) yields about a 20 per cent reduction. For income, \( M_5 \) yields about a 3 per cent reduction in RSS relative to \( M_1 \) while \( M_7 \) provides the lowest RSS, about 6 per cent lower than for \( M_1 \).

For nested models, a large sample likelihood ratio test procedure has been employed to compare alternative formulations with results reported in table 9.
For consumption, $M_2$ is preferred to $M_1$ at the 5 per cent level. However, pairwise comparisons of $M_2$ against $M_4$, $M_5$ and $M_6$ favor the latter relative to $M_2$. However, in comparisons with $M_7$, both $M_2$ and $M_4$ are rejected. Thus the results of the likelihood ratio tests favor $M_7$, a model that is compatible with the results of the final equation analysis under the assumption $h_{23}(L) \equiv h_{33}(L)$. The results for income transfer functions, shown in the bottom of Table 9, indicate that $M_2$ is rejected in favor of $M_1$ while $M_4$ performs better than $M_3$ or $M_6$. Compared with $M_7$, models $M_2$, $M_4$ and $M_6$ are rejected at the 5 per cent significance level, it appears that $M_7$ is not significantly different from $M_5$. The results of these comparisons suggest that it is reasonable to accept tentatively, models $M_1$, $M_5$ or $M_7$ as being in accord with the information in the data. If we retain models $M_7$ for consumption and $M_7$ for income, we have $r_{12} + r_{21} = 5$ for the order of the AR polynomials acting on $\Delta c_t$ and $\Delta y_t$, the degrees of the polynomials operating on $\Delta x_t$ are of degrees, 5 and 4, respectively, and the error processes are each of order 4. Under the assumption that $h_{23}(L) \equiv h_{33}(L)$, these results are in accord with the requirements that the final equations must satisfy (see Table 2).

6. Summary of results and implications for structural equations

In Table 10, we present the preferred final equation and transfer function models for the dynamized Haavelmo model. From the information provided in Table 10, the following are the implied restrictions on the lag structures appearing in the structural equations of the model where the $r_{ij}$'s refer to the degrees of elements of $H(L)$, the matrix $H(L)$ divided by $(1-L)$:

1. $r_{12} = 1$; $r_{33} = r_{23} = r_{21} = 4$; $q_{11} = q_{12} = 0$; and $q_{21}, q_{22} \leq 3$, with at least one equality holding.

2. The transfer functions show a dependence of $\Delta c_t$ and of $\Delta y_t$ on $\Delta x_t$. Under the assumption that $h_{23} = h_{33}$, the final equations and transfer functions selected by the likelihood ratio tests are compatible insofar as the degrees of the relevant lag polynomials are considered.

3. Explicitly, a structural representation compatible with the results of the

---

18 Other possibilities, e.g. $M_7$ for $\Delta c_t$ and $M_1$ for $\Delta y_t$ or $M_7$ for $\Delta c_t$ and $M_4$ for $\Delta y_t$, lead to incompatibilities with the requirements that the final and transfer equations must satisfy. In the first case, $M_7$ for $\Delta c_t$ and $M_1$ for $\Delta y_t$, the AR parts of the transfer functions are not identical as required in (3.14)–(3.15) and even possible cancelling will not be sufficient to satisfy the condition on the polynomials hitting $\Delta x_t$ in (3.14)–(3.15). If we retain $M_7$ for $\Delta c_t$ and $M_4$ for $\Delta y_t$, their autoregressive parts have the same order, $r_{12} + r_{21} = 5$, and the degrees of the polynomials for $\Delta x_t$ are respectively $r_{12} + r_{23} = 5$ and $r_{23} = 3$, implying $r_{21} = 2$. However, the assumption $h_{23}(L) \equiv h_{33}(L)$, implying $r_{21} = r_{23}$, is no longer satisfied. In addition, there is incompatibility with the analysis of the final equations requiring $r_{23} = 4$.
Table 10
Final equation and transfer function models for dynamized Haavelmo model.*

<table>
<thead>
<tr>
<th>Systems of equations</th>
<th>Model</th>
<th>Order of AR part</th>
<th>Degree of lag polynomial for $\Delta x_t$</th>
<th>Order of MA error process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Final equations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta c_t$</td>
<td>(5, 1, 4)*</td>
<td>$r_{12} + r_{21} = 5$</td>
<td>$q_{11}, q_{12}, r_{12} + q_{21}, r_{12} + q_{22} \leq 4$ (at least one equality holding)</td>
<td></td>
</tr>
<tr>
<td>$\Delta y_t$</td>
<td>(5, 1, 4)*</td>
<td>$r_{12} + r_{21} = 5$</td>
<td>$q_{21}, q_{22}, r_{21} + q_{22}, r_{21} + q_{12} \leq 4$ (at least one equality holding)</td>
<td></td>
</tr>
<tr>
<td>$\Delta x_t$</td>
<td>(4, 1, 0)*</td>
<td>$r_{32} = 4$</td>
<td>$q_{33} = 0$</td>
<td></td>
</tr>
<tr>
<td>2. Transfer functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta c_t$</td>
<td>$M_7^e$</td>
<td>$r_{12} + r_{21} = 5$</td>
<td>$r_{12} + r_{23} = 5$</td>
<td>$q_{11}, q_{12}, r_{12} + q_{21}, r_{12} + q_{22} \leq 4$ (at least one equality)</td>
</tr>
<tr>
<td>$\Delta y_t$</td>
<td>$M_7^e$</td>
<td>$r_{12} + r_{21} = 5$</td>
<td>$r_{33} = 4$</td>
<td>$q_{21}, q_{22}, r_{12} + r_{21}, q_{12} + r_{21} \leq 4$ (at least one equality)</td>
</tr>
</tbody>
</table>

*It is assumed that $h_{23} = h_{33}$.

*See tables 3–5 where estimated models are presented.

*See table 8 where estimated models are presented.

The final equation and transfer function analyses is:

$$
\begin{bmatrix}
1 - \alpha^{(1)} \\
\mu_1^{(4)} - 1 & 1 & \mu_2^{(4)} - 1 \\
0 & 0 & \mu_2^{(4)} - 1
\end{bmatrix}
\begin{bmatrix}
c_{1t} \\
y_{1t} \\
x_{1t}
\end{bmatrix}
= 
\begin{bmatrix}
\theta_1 \\
\theta_2 \\
\theta_3
\end{bmatrix}
$$

where the superscripts in parentheses denote the degrees of lag polynomials that were determined from the final equation and transfer function analyses. Note that these polynomials are equal to the polynomials of the matrix $H(L)$ in (3.8) divided by a common factor $1 - L$. The factor $1 - L$ hitting the variables $c_t$, $y_t$, and $x_t$ puts them in first difference form, a transformation that appears adequate.
to induce stationarity in all three variables, a condition required for the correlogram analysis of the variables. That the same differencing transformation induces stationarity in all variables is not necessary for all models but is an empirical finding in the present case. Also, to achieve compatibility, it is necessary that \( h_{11}(L) = h_{22}(L) = 1 \), a special case of what was assumed in (3.7a). Last, it should be noted that \( \mu_1^{(4)}(L) \) and \( \mu_2^{(4)}(L) \) are not necessarily identical.

The system in (6.1) can alternatively be expressed in the form of (3.2a)–(3.2c) as follows:

\[
\begin{bmatrix}
1 & 0 & -\alpha^{(1)} & 0 \\
-\mu_1^{(4)} & 1 & 0 & -\mu_2^{(4)} \\
-1 & 1 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
c_t \\
r_t \\
y_t \\
x_t
\end{bmatrix}
= (1-L)
\begin{bmatrix}
\beta \\
v \\
0
\end{bmatrix}
+ \begin{bmatrix}
u_t \\
w_t \\
0
\end{bmatrix},
\]

with \( u_t \) a serially uncorrelated disturbance term and \( w_t \) following a third order moving average process. Further, \( u_t \) and \( w_t \) will generally be correlated.

Using the identity, \( \Delta y_t = \Delta c_t + \Delta x_t - \Delta r_t \), we can eliminate \( \Delta r_t \) from (6.2) to obtain:

\[\Delta c_t = \alpha_0 \Delta y_t + \alpha_1 \Delta y_{t-1} + \beta + u_t, \tag{6.3a}\]

and

\[\Delta y_t = (1 - \mu_1^{(4)}) \Delta c_t + (1 - \mu_2^{(4)}) \Delta x_t - v - w_t, \tag{6.3b}\]

where \( \alpha^{(1)} \equiv \alpha_0 + \alpha_1 L, \ 1 - \mu_1^{(4)} \equiv \Sigma_{i=0}^{4} \gamma_i L^i, \) and \( -(1 - \mu_2^{(4)}) \Delta x_t = f_{30}^{(0)} e_3 \), have been used and \( w_t' \equiv -(w_t + f_{33}^{(0)} e_3) \). The two equation system in (6.3) is a simultaneous equation model with dynamic lags and contemporaneously correlated disturbance terms, \( u_t \) and \( w_t' \), the former non-autocorrelated and the latter following a third order MA process. We can estimate the parameters of (6.3) employing 'single equation' or 'joint' estimation techniques as explained briefly below.\(^{19}\)

For single equation estimation of (6.3a), we consider it in conjunction with the final equation for \( y_t \), namely a \((5, 1, 4)\) ARMA process that we write as:

\[\Delta y_t = \sum_{i=1}^{5} \delta_i \Delta y_{t-i} + \phi_1 + \sum_{i=0}^{4} \lambda_{1i} q_{1t-i}, \tag{6.4}\]

where \( a_{1t} \) is a non-autocorrelated error with zero mean and constant finite

\(^{19}\)These estimation procedures will be treated more fully in future work. A recent paper by Byron (1973) treats some of these problems from the likelihood point of view. Also, it will be noted that non-unique estimates for certain parameters are available from the final equation and transfer function analyses. In certain instances these latter estimates are obtained from estimates of ratios of lag polynomials and thus are probably not very reliable.

\(^{20}\)Alternatively, the transfer function for \( \Delta y_t \) could be employed. However, it is not clear that use of the transfer function is to be preferred.
variance. The parameters of (6.4) have already been estimated above. We now substitute for Δy_i in (6.3a) from (6.4) to obtain

$$Δc_t = θ_0 Δy_t + θ_1 Δy_{t-1} + β' + v_{1t},$$

(6.5)

where $Δy_t = ∑_{i=1}^{4} θ_i Δy_{t-i}$, $β' = θ(θ_0 + θ_1)$ and $v_{1t} = w_t + θ_0 ∑_{i=1}^{4} θ_i a_{1t-i}$ a fourth order MA process. Given consistent estimates of the $θ_i$'s in $Δy_t$, we can calculate consistent estimates of $θ_0$, $θ_1$, $θ$ and parameters of the MA process for $v_{1t}$. The results of this approach are presented and discussed below.

With respect to single equation estimation of (6.3b), we consider it in conjunction with the (5, 1, 4) ARMA final equation for $c_t$ that was estimated above and is expressed as:

$$Δc_t = ∑_{i=1}^{4} π_i Δc_{t-i} + φ_2 + ∑_{i=0}^{4} λ_2 a_{2t-i},$$

(6.6)

where $a_{2t}$ is a non-autocorrelated error with zero mean and constant finite variance. Then on substituting for $Δc_t$ in the second line of (6.3b) from (6.6), we have:

$$Δy_t = γ_0 Δc_t + ∑_{i=1}^{4} γ_i Δc_{t-i} + ν' + v_{2t},$$

(6.7)

where $Δc_t = ∑_{i=1}^{4} π_i Δc_{t-i}$, $ν' = ν + γ_0 φ_2$, and $v_{2t} = w_t + γ_0 ∑_{i=1}^{4} λ_2 a_{2t-i}$, a fourth order MA process. Since consistent estimates of the $π_i$ are available from the analysis of (6.6), they can be used in conjunction with (6.7) to obtain consistent estimates of the $θ'$s, $ν$ and the parameters of the process for $v_{2t}$.

As regards joint estimation of (6.5) and (6.7), single equation analysis yields residuals that can be used to estimate the covariance matrix for the disturbances, the $v_{1t}$'s and $v_{2t}$'s. For a two equation system, this matrix will be generally a $2T × 2T$ matrix with four submatrices in the form of band matrices characteristic of MA processes. Let this matrix be denoted $Ω$ and an estimate of it, $Ω̂$. Then with $v' = (v_1' v_2')$, where the vector $v_1$ has elements $v_{1t}$ and $v_2$ elements $v_{2t}$, minimization of $v'Ω̂^{-1}v$ can be done to provide joint estimates of the parameters.21

In table 11, we present various single equation consistent estimates of the parameters of the consumption equation in (6.3a). In the first line of the table, the final equation for $Δy_t$ was employed to substitute for $Δy_t$ in the consumption function while in the second line the transfer function for $Δy_t$ was employed.22

It is seen that in both cases the point estimate for $θ_0$ is negative. However, the standard errors are large so that a confidence interval at a reasonable level would include positive values. The estimates of $θ_1$, the coefficient of $Δy_{t-1}$ in (6.3a) are

21The new residuals can be employed to reestimate $Ω$ and thus iteration of the process on $Ω$ (and also on the parameters in $Δc_t$ and $Δy_t$) is possible.

22Note that the estimation of the consumption equation using the final equation expression for $Δy_t$ is not linked to the assumption that $Δx_t$ is exogenous whereas use of the transfer function expression for $Δy_t$ is.
A. Zellner, F. Palm, *Time series and econometric models*

in the vicinity of 0.3 with a standard error of about 0.1.\(^3\) That \(x_0\) and \(x_1\) are not very precisely estimated is probably due to collinearity of \(\Delta y_t\) and \(\Delta y_{t-1}\). Use of an informative prior distribution for \(x_0\) and \(x_1\) in a Bayesian analysis could help to improve the precision of inferences. To specify a prior distribution for \(x_0\) and \(x_1\) and also to interpret the results in table 11, it may be useful to regard \(\Delta c_t\), the planned change in expenditures, including durables, to be linked to permanent income change, \(\Delta y^p_t\), and transitory income change, \(\Delta y^t_t\), as follows: \(\Delta c_t = k\Delta y^p_t + x_0\Delta y^t_t + x_1\Delta y^t_{t-1}\). In planning consumption expenditures for the \(r\)th period, note that \(\Delta y^t_t\) is as yet unrealized transitory income whereas \(\Delta y^t_{t-1}\) is realized transitory income for period \(t-1\). We believe that consumer reactions to realized transitory income will be much greater than that to as yet unrealized transitory income, i.e. \(x_1 > x_0\) with \(x_0\) small. Using \(\Delta c_r = \Delta c^p_r + u_r\) and \(\Delta y_r = \Delta y^p_r + \Delta y^t_r\) in connection with the relation for \(\Delta c^p_r\) above, we have \(\Delta c_r = x_0\Delta y^p_r + x_1\Delta y^t_{r-1} + \beta + u_r\) with \(\beta = (k-x_0)\Delta y^p_r - x_1\Delta y^p_{r-1}\), assumed constant.\(^4\) Within this framework, given the hypothesis that reaction to unrealized transitory income change, \(\Delta y^t_r\) is rather small, if not zero, while reaction to realized transitory income change \(\Delta y^t_{r-1}\) is positive, probably an \(x_1\) between zero and one, the results in table 11 appear plausible.

In conclusion, we believe that the techniques presented above can be very helpful in checking the specifying assumptions of many existing linear or linearized models and in ‘iterating in’ on models that are suitable approximations to

\(^3\)As explained below \(x_1\) can be viewed as the coefficient of realized transitory income change and thus an estimate of \(x_1\) in the vicinity of 0.3 seems reasonable.

\(^4\)Alternatively, we could assume \((k-x_0)\Delta y^p_r - x_1\Delta y^p_{r-1} = \beta + \epsilon_r\), where \(\epsilon_r\) is a non-autocorrelated random error with zero mean and constant variance.
the information in our data and that may predict well. Some topics that will receive attention in future work include further development of estimation techniques for different equation systems, joint testing procedures for nested and non-nested hypotheses, analyses of the comparative predictive performance of final equation, transfer function and structural equation systems, Bayesian procedures utilizing informative prior distributions, and applications. Finally, we cannot resist remarking that the present work lends support to the notion that so-called 'naive' ARMA time series models are not all that naive after all.

References


Byron, R.P., 1973, The computation of maximum likelihood estimates for linear simultaneous systems with moving average disturbances (Department of Economics, Australian National University) 19 pp. manuscript.


Nelson, C.R., 1970, Joint estimation of parameters of correlated time series (Graduate School of Business, University of Chicago) 41 pp. manuscript.

Palm, F., 1972, On mixed prior distributions and their application in distributed lag models, CORE Discussion Paper 7222 (University of Louvain).


After completing this paper, the following Ph.D. Thesis, dealing with related topics, was brought to our attention by Dennis Aigner:

Sources of data

Personal consumption expenditures, disposable personal income, gross investment data
Series 1946–65:
Series 1966–72:
Consumer price index:
Population data: